

Two-way quantum communication channels

ANDREW M. CHILDS

*Institute for Quantum Information
California Institute of Technology
Pasadena, CA 91125, USA
amchilds@caltech.edu*

DEBBIE W. LEUNG

*Institute for Quantum Information
California Institute of Technology
Pasadena, CA 91125, USA
wcleung@cs.caltech.edu*

HOI-KWONG LO

*Center for Quantum Information and Quantum Control,
Department of Electrical & Computer Engineering,
and Department of Physics
University of Toronto
Toronto, Ontario M5S 3G4, Canada
hklo@comm.utoronto.ca*

We consider communication between two parties using a bipartite quantum operation, which constitutes the most general quantum mechanical model of two-party communication. We primarily focus on the simultaneous forward and backward communication of classical messages. For the case in which the two parties share unlimited prior entanglement, we give inner and outer bounds on the achievable rate region that generalize classical results due to Shannon. In particular, using a protocol of Bennett, Harrow, Leung, and Smolin, we give a one-shot expression in terms of the Holevo information for the entanglement-assisted one-way capacity of a two-way quantum channel. As applications, we rederive two known additivity results for one-way channel capacities: the entanglement-assisted capacity of a general one-way channel, and the unassisted capacity of an entanglement-breaking one-way channel.

1. Introduction

Suppose two parties, Alice and Bob, wish to exchange information. To do so, they must be connected by some physical interaction, or in information-theoretic language, a channel. One of the main problems of information theory is to determine the maximum rates (i.e., the capacities) for communication through such a channel.

A particularly well-studied type of interaction is the *one-way channel*, which models transporting some carrier of information from a fixed sender (say, Alice) to a fixed receiver (say, Bob), with the state encoding the information possibly

modified in some way during transit. In other words, a one-way channel is a formal change of ownership of the state together with a state change.

Classical one-way channels, introduced by Shannon,¹ are stochastic maps on probability distributions. More generally, in quantum mechanics, evolution is described by a quantum operation, i.e., a trace-preserving, completely positive map on quantum states represented by density matrices (positive semidefinite matrices of unit trace). Figure 1(a) shows a one-way quantum channel described by the quantum operation \mathcal{M} . Alice prepares an input state ρ_A which is transported to Bob as the output state $\rho_B = \mathcal{M}(\rho_A)$. The capacities to transmit classical²⁻⁷ and quantum⁸⁻¹¹ messages through such channels have been extensively studied.

Although a one-way channel is an intuitive model for communication, it is only a special case of the possible interactions between Alice and Bob. The most general interaction is a joint quantum operation \mathcal{N} acting on the joint Hilbert space of the two parties. If they are in possession of a joint density matrix ρ_{AB} , then the action of the channel produces a joint output state $\rho'_{AB} = \mathcal{N}(\rho_{AB})$. In other words, Alice and Bob each provide an input to the two-way channel, which evolves their inputs jointly to produce a joint output shared by the two parties. Such a *two-way quantum channel* is shown in Fig. 1(b).

The two-way quantum channel model is the most general setting for two-party communication. For example, one-way channels are simply two-way channels with a zero-dimensional input for Bob and a zero-dimensional output for Alice (or equivalently from an operational standpoint, channels that discard Bob's input and give Alice a fixed output).

Another subclass of two-way quantum channels consists of the classical two-way channels. Such channels were also first studied by Shannon,¹² who gave outer and inner bounds on their capacity regions. Shannon's bounds were subsequently

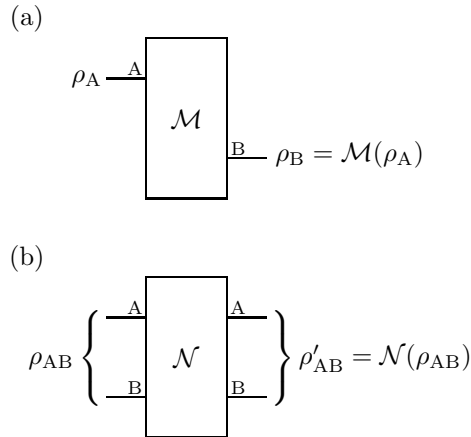


Fig. 1. (a) One-way and (b) two-way quantum channels.

improved by many others (see for example Refs. 13–20).

In the quantum setting, another natural special class of two-way channels is the set of bipartite unitary interactions acting on systems of fixed dimension. The capacity question in this setting was formalized and studied in Ref. 21, and further capacity expressions were subsequently found in Refs. 22 and 23. For such channels, the problem of bidirectional communication is closely related to the problem of generating entanglement.^{21–25} Generalizing to allow systems whose input and output dimensions are different, one finds an especially simple class of interactions, the *quantum feedback channels*.²⁶ Such channels take no input from Bob and evolve Alice’s input into a state shared between her and Bob.

A unifying viewpoint is that any two-way quantum channel can be viewed as an isometry of the two input states to three output states, discarding the third part of the system to an inaccessible environment. A variety of subclasses can be obtained simply by changing the dimensions of the terminals.

The various subclasses of two-way channels can exhibit remarkably different properties. For example, any (nonlocal) two-way unitary channel can communicate in either direction and generate entanglement at a nonzero rate,²¹ while in each of the other examples, some of these tasks are known to be impossible. Because of the wide variety of possible subclasses of two-way channels, and because calculating capacities is known to be difficult for several of the possibilities, we do not expect especially tight capacity results for the general two-way channel.

A communication protocol using two-way channels may yield classical or quantum communication in either or both directions. It can also create or consume classical or quantum correlations as auxiliary resources. In particular, providing Alice and Bob with enough free entanglement considerably simplifies and unifies the study of communication capacities of one-way quantum channels^{6,7} as well as two-way channels.^{21–23,25}

In this paper, we will mostly be considering the set of achievable rates of classical communication R_{\Rightarrow} from Alice to Bob and R_{\Leftarrow} from Bob to Alice, and the rate of producing entanglement R_e (which can either be positive, indicating that entanglement is produced, or negative, indicating that entanglement is consumed). The set of achievable rates $(R_{\Rightarrow}, R_{\Leftarrow}, R_e)$ forms a three-dimensional region. The boundary of this region represents the set of best achievable rates. Figure 2 shows a schematic diagram of the achievable rates of classical communication $(R_{\Rightarrow}, R_{\Leftarrow})$ at fixed R_e . Typically, there is a tradeoff between how much information Alice can send to Bob and how much information Bob can send to Alice. The end points of the optimal tradeoff curve, where $R_{\Leftarrow} = 0$ or $R_{\Rightarrow} = 0$, are called the *one-way capacities* of the two-way channel. The cases in which $R_e = 0$ and $R_e \rightarrow -\infty$, called *entanglement-unassisted* (or simply *unassisted*) and *entanglement-assisted*, respectively, are of particular interest.

For the case of entanglement-assisted one-way classical communication by unitary two-way channels, Ref. 21 gave a simplified single-letter expression for the

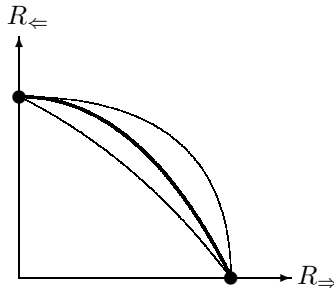


Fig. 2. The achievable region for two-way classical communication at fixed R_e , with outer and inner bounds.

capacity, showed that it is additive, and gave a protocol for achieving it. These results were subsequently extended in Ref. 22 to the case in which a fixed amount of entanglement is consumed or generated (a scenario dubbed *finite entanglement assistance*). In addition, tradeoff curves $\{(R_{\Rightarrow}, 0, R_e)\}$ for one-way classical communication were related to analogous curves $\{(R_{\Leftarrow}, 0, R_e)\}$ for quantum communication. In Ref. 23, the entire three-dimensional achievable regions for finite entanglement-assisted bidirectional classical or quantum communication were related.

In this paper, we consider bidirectional classical communication using a two-way quantum channel in the general (not necessarily unitary) case, possibly allowing entanglement assistance. We begin in Sec. 2 by introducing some basic notation and formalizing the notions of a protocol and the capacities it achieves. In Sec. 3, we extend techniques from Ref. 21 to provide inner and outer bounds on the achievable region. The outer bound generalizes Shannon's bounds¹² to the quantum setting. The inner bound is an extension of the protocol of Ref. 21 for simultaneous two-way communication, and also reduces to Shannon's bound in the classical case. Furthermore, the bounds meet on the axes (as depicted in Fig. 2), giving a formula for the one-way capacity of a two-way quantum channel (extending the unitary result of Ref. 21). In Sec. 4, we describe two immediate applications of the inner and outer bounds. First, as a simple consequence of the fact that the inner and outer bounds meet on the axes, we recover the additivity of the entanglement-assisted classical capacity of one-way channels.⁷ We also provide a simple operational derivation of the additivity of the unassisted classical capacity of an entanglement-breaking channel, a special case of a result that was first proved in Ref. 27. Finally, in Sec. 5, we discuss some open questions and directions for future investigation.

2. Preliminaries

In this section, we describe our framework and define the notation used throughout the paper.

An *ebit* refers to a unit of shared quantum correlation, as quantified by an EPR pair of qubits $|\Phi\rangle_{AB} := \frac{1}{\sqrt{2}} \sum_{x=0}^1 |x\rangle_A |x\rangle_B$, with the density matrix $\Phi := |\Phi\rangle\langle\Phi|$.

Throughout the paper, we omit the tensor product symbol, \otimes , if no confusion will arise. The functions \exp and \log are always base 2. The positive and negative parts of a real number x are written as x^\pm : for $x \geq 0$, $x^+ := x$ and $x^- := 0$; for $x < 0$, $x^+ := 0$ and $x^- := -x$. We adopt the convention that a quantum operation \mathcal{N} acts on an ensemble of quantum states $\mathcal{E} = \{p_i, \rho_i\}$ by acting on each state individually (preserving its probability), i.e., $\mathcal{N}\mathcal{E} = \{p_i, \mathcal{N}(\rho_i)\}$.

Our framework for communication using two-way channels extends the model of Refs. 21 and 23. A general two-way quantum channel \mathcal{N} has two inputs of dimensions $d_A^{\text{in}}, d_B^{\text{in}}$ and two outputs of dimensions $d_A^{\text{out}}, d_B^{\text{out}}$. A general protocol \mathcal{P}_n for two-way classical communication with n uses of \mathcal{N} consists of n alternating steps of local processing and application of \mathcal{N} followed by a final step of local processing, as depicted in Fig. 3. Let a, b be the respective messages to be communicated from Alice to Bob and vice versa, and let $\mathcal{A}_k, \mathcal{B}_k$ be the local operations of Alice and Bob between the k th and $(k+1)$ th use of \mathcal{N} . The operations $\mathcal{A}_k, \mathcal{B}_k$ may be arbitrary, but without loss of generality, we can assume that all ancillas are present at the beginning of the protocol and that all measurements are performed coherently, with no systems discarded by $\mathcal{A}_k, \mathcal{B}_k$. In other words, all of $\mathcal{A}_k, \mathcal{B}_k$ can be assumed to be unitary. Alice's initial operation \mathcal{A}_0 has three input systems: A_m , which contains the message a ; A_a , which contains a local ancilla; and A_e , which is maximally entangled with B_e . The operation \mathcal{A}_0 converts $A_{m,a,e}$ unitarily into two systems: A , which is the input to \mathcal{N} , and A' , which is not acted on by \mathcal{N} . Each of \mathcal{A}_k for $k = 1, \dots, n-1$ has two inputs A, A' and two outputs, which for simplicity will also be labeled as A, A' , although in general they may have different dimensions than the input systems. The operation \mathcal{A}_k changes the dimensions of A, A' if $d_A^{\text{in}} \neq d_A^{\text{out}}$. Finally, \mathcal{A}_n converts A, A' to three output systems: A_m , whose reduced state represents Bob's message for Alice; A_a , which is an arbitrary ancillary system; and A_e , which is nearly maximally entangled with B_e . The situation is analogous on Bob's side.

The goal is to find a family of protocols $\{\mathcal{P}_n : n \in \mathbb{N}\}$, where protocol \mathcal{P}_n employs n uses of \mathcal{N} , that will perform the communication task with accuracy that can be made arbitrarily good by taking n sufficiently large—the output state

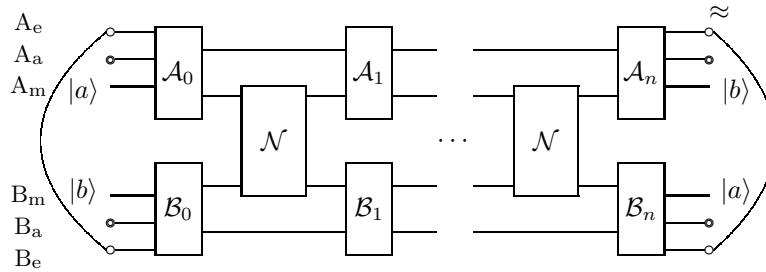


Fig. 3. Circuit representation of a general two-way channel protocol.

of A_m, B_m should be close to $|b\rangle\langle b| \otimes |a\rangle\langle a|$, the state of A_e, B_e should be nearly maximally entangled, and A_a, B_a should be nearly disentangled from $A_{m,e}, B_{m,e}$ (but can depend on the messages a, b). The input dimensions of A_m, B_m determine the communication rates, while the change of the dimensions of A_e, B_e determines the amount of entanglement consumed or generated by \mathcal{P}_n .

To quantify the proximity of two quantum states ρ, σ , we use the trace distance $\frac{1}{2} \|\rho - \sigma\|_1$, where $\|X\|_1 := \text{tr} \sqrt{X^\dagger X}$. For example, for two pure states $|\alpha\rangle, |\beta\rangle$, $\frac{1}{2} \|\alpha\rangle\langle\alpha| - |\beta\rangle\langle\beta|\|_1 = \epsilon \Leftrightarrow |\langle\alpha|\beta\rangle|^2 = 1 - \epsilon^2$.

We are now ready to define the achievable region:

Definition 1 (Achievable rates). We say that $(R_{\Rightarrow}, R_{\Leftarrow}, R_e)$ is achievable if there is a sequence of protocols $\{\mathcal{P}_n\}$, together with asymptotically vanishing sequences $\{\delta_n\}, \{\epsilon_n\}$ of nonnegative numbers and sequences of integers $\{r_{\Rightarrow,n}\}, \{r_{\Leftarrow,n}\}, \{r_{e,n}\}$ such that

$$r_{\Rightarrow,n} \geq n(R_{\Rightarrow} - \delta_n), \quad (1)$$

$$r_{\Leftarrow,n} \geq n(R_{\Leftarrow} - \delta_n), \quad (2)$$

$$r_{e,n}^+ \geq n(R_e^+ - \delta_n), \quad (3)$$

$$r_{e,n}^- \leq n(R_e^- + \delta_n), \quad (4)$$

and for all messages $a \in \{0, 1\}^{r_{\Rightarrow,n}}, b \in \{0, 1\}^{r_{\Leftarrow,n}}$ the following success criteria hold. Let

$$\rho_n^{ab} := \mathcal{P}_n(|0\rangle\langle 0|_{A_a}|0\rangle\langle 0|_{B_a} \otimes |a\rangle\langle a|_{A_m}|b\rangle\langle b|_{B_m} \otimes \Phi_{A_e B_e}^{\otimes r_{e,n}^-}), \quad (5)$$

and let $\rho_{n, A_m B_m}^{ab}, \rho_{n, A_e B_e}^{ab}$ be its reductions to the message and entanglement subspaces, respectively. Then the success criteria are

$$\frac{1}{2} \|\rho_{n, A_m B_m}^{ab} - |b\rangle\langle b|_{A_m}|a\rangle\langle a|_{B_m}\|_1 \leq \epsilon_n, \quad (6)$$

$$\frac{1}{2} \left\| \rho_{n, A_e B_e}^{ab} - \Phi_{A_e B_e}^{\otimes r_{e,n}^+} \right\|_1 \leq \epsilon_n. \quad (7)$$

The achievable region is *convex* by convex combination (i.e., time sharing) of protocols. It is *monotone* since resources can always be discarded to give a lower rate of communication, lower rate of entanglement production, or higher rate of entanglement consumption.

Any quantum protocol for communicating classical information concludes by decoding the quantum state to produce a classical output. The final decoding can be viewed as a measurement of the final state, which in our model is implemented unitarily. The general problem of extracting classical information encoded in a quantum state (i.e., learning the index i in a random draw from an ensemble $\mathcal{E} = \{p_i, \rho_i\}$ of quantum states) has been thoroughly studied in this context.²⁻⁵ The accessible information of \mathcal{E} , denoted $I_{\text{acc}}(\mathcal{E})$, is defined to be the maximum mutual information between the label i and the outcome of any possible measurement. It is upper bounded as²

$$I_{\text{acc}}(\mathcal{E}) \leq \chi(\mathcal{E}), \quad (8)$$

where $\chi(\mathcal{E})$ denotes the *Holevo information* of \mathcal{E} ,

$$\chi(\mathcal{E}) := S\left(\sum_i p_i \rho_i\right) - \sum_i p_i S(\rho_i). \quad (9)$$

Here $S(\rho) := -\text{tr } \rho \log \rho$ is the von Neumann entropy of ρ .

In fact, by appropriate encoding, the Holevo information turns out to be asymptotically achievable in the following sense. Consider a *CQ channel*, i.e., a one-way channel with a classical input i giving rise to a corresponding quantum output state ρ_i . Then we have

Theorem 1 (HSW Theorem^{3–5}). The capacity of the CQ channel $\{i \rightarrow \rho_i\}$ is given by $\sup_{\{p_i\}} \chi(\{p_i, \rho_i\})$.

More specifically, for any fixed input probability distribution $\{p_i\}$, let $\mathcal{E} = \{p_i, \rho_i\}$ denote the corresponding ensemble. In the limit of large n , there is a set of $\exp(n\chi(\mathcal{E}) - \delta_n)$ states $\rho_{i_1} \otimes \rho_{i_2} \otimes \cdots \otimes \rho_{i_n}$ that can be distinguished with error bounded by some ϵ_n , where $\delta_n, \epsilon_n \rightarrow 0$ as $n \rightarrow \infty$. This is proved by showing that, with nonzero probability, a random code (in which each tensor component of each codeword is drawn independently from \mathcal{E}) can be decoded with vanishing error.

The HSW theorem can be applied to an arbitrary one-way channel by viewing the states ρ_i as the possible outputs of the channel. However, for a general two-way channel, the coding problem is complicated somewhat by the fact that the input and output ensembles are bipartite. In this case, communication in one direction is typically affected by the input to the channel from the opposite direction. Thus, we will need to prove a modified version of the HSW theorem when we derive protocols for bidirectional communication.

3. Entanglement-assisted capacity of two-way quantum channels

In this section, we present inner and outer bounds on the classical capacity region of an entanglement-assisted two-way quantum channel.

Let \mathcal{P}_n be an arbitrary protocol that employs n uses of the two-way channel \mathcal{N} . We assume without loss of generality that Alice and Bob retain a copy of their input messages throughout the protocol. Then, after t uses of \mathcal{N} followed by $\mathcal{A}_t \otimes \mathcal{B}_t$, Alice and Bob possess a joint state drawn from an ensemble

$$\mathcal{E}^{(t)} = \{p_{ab}, |a\rangle\langle a| \otimes \rho_{ab}^{(t)} \otimes |b\rangle\langle b|\} \quad (10)$$

indexed by the classical messages a, b to be communicated, where $\mathcal{E}^{(t)}$ is completely determined by \mathcal{P}_n .

For a bipartite ensemble \mathcal{E} over systems AA' and BB' , we define the *local Holevo information* of BB' and AA' , respectively, as

$$\chi_{\Rightarrow}(\mathcal{E}) := \chi(\text{tr}_{AA'} \mathcal{E}), \quad (11)$$

$$\chi_{\Leftarrow}(\mathcal{E}) := \chi(\text{tr}_{BB'} \mathcal{E}). \quad (12)$$

Following earlier work,²¹ our analysis will be based on examining the local Holevo information $\chi_{\Rightarrow}(\mathcal{E}^{(t)})$, $\chi_{\Leftarrow}(\mathcal{E}^{(t)})$ for $t = 0, 1, \dots, n$ during the course of an n -use protocol. By the joint entropy theorem (see for example Eq. (1.58) in Ref. 28), for an ensemble $\mathcal{E}^{(t)}$ of the form of Eq. (10) (i.e., with local copies of the classical messages), the local Holevo information can be rewritten as

$$\chi_{\Rightarrow}(\mathcal{E}^{(t)}) = H(\{p_b\}) + \sum_b p_b \chi(\mathcal{E}_b^{(t)}), \quad (13)$$

$$\chi_{\Leftarrow}(\mathcal{E}^{(t)}) = H(\{p_a\}) + \sum_a p_a \chi(\mathcal{E}_a^{(t)}), \quad (14)$$

where $p_a := \sum_b p_{ab}$ and $p_b := \sum_a p_{ab}$ are the marginal distributions and $\mathcal{E}_b^{(t)} := \{p_{a|b}, \text{tr}_{AA'} \rho_{ab}^{(t)}\}$, $\mathcal{E}_a^{(t)} := \{p_{b|a}, \text{tr}_{BB'} \rho_{ab}^{(t)}\}$ are the ensembles for Bob and Alice conditioned on their known inputs b and a , respectively, where $p_{\cdot|\cdot}$ denotes conditional probability.

Equations (13) and (14) give natural interpretations of the local Holevo information: for example, for Bob, it is the sum of the information about b that he already knows and the information about a obtainable from $\mathcal{E}_b^{(t)}$, averaged over b . Removing the information that the sender already knows gives a quantity that is useful for obtaining bounds. Thus, we define the *readjusted local Holevo information* as

$$\bar{\chi}_{\Rightarrow}(\mathcal{E}^{(t)}) := \sum_b p_b \chi(\mathcal{E}_b^{(t)}), \quad (15)$$

$$\bar{\chi}_{\Leftarrow}(\mathcal{E}^{(t)}) := \sum_a p_a \chi(\mathcal{E}_a^{(t)}). \quad (16)$$

3.1. Additive outer bound

In this section, we obtain an additive outer bound on the capacity region for entanglement-assisted bidirectional classical communication using a two-way quantum channel.

Consider the differences in local Holevo information induced by an application of \mathcal{N} to an ensemble \mathcal{E} :

$$\Delta\chi_{\Rightarrow}(\mathcal{E}) := \chi_{\Rightarrow}(\mathcal{N}\mathcal{E}) - \chi_{\Rightarrow}(\mathcal{E}) = \bar{\chi}_{\Rightarrow}(\mathcal{N}\mathcal{E}) - \bar{\chi}_{\Rightarrow}(\mathcal{E}), \quad (17)$$

$$\Delta\chi_{\Leftarrow}(\mathcal{E}) := \chi_{\Leftarrow}(\mathcal{N}\mathcal{E}) - \chi_{\Leftarrow}(\mathcal{E}) = \bar{\chi}_{\Leftarrow}(\mathcal{N}\mathcal{E}) - \bar{\chi}_{\Leftarrow}(\mathcal{E}). \quad (18)$$

Let $\llbracket x, y \rrbracket$ denote the region $[0, x] \times [0, y] \subset \mathbb{R}^2$, and let $\text{conv}(\cdot)$ denote the convex hull. In terms of these quantities, we have the following outer bound on the achievable region:

Theorem 2. If $(R_{\Rightarrow}, R_{\Leftarrow}, -\infty)$ is achievable, then

$$(R_{\Rightarrow}, R_{\Leftarrow}) \in \text{conv}\{\llbracket \Delta\chi_{\Rightarrow}(\mathcal{E}), \Delta\chi_{\Leftarrow}(\mathcal{E}) \rrbracket : \text{arbitrary } \mathcal{E}\}. \quad (19)$$

Proof. For any \mathcal{P}_n , the Holevo bound on the accessible information, Eq. (8), implies that the number of bits that can be faithfully transmitted forward and backward are no more than $\bar{\chi}_{\Rightarrow}(\mathcal{E}^{(n)})$ and $\bar{\chi}_{\Leftarrow}(\mathcal{E}^{(n)})$ respectively. In other words,

$$R_{\Rightarrow} \leq \frac{1}{n} \bar{\chi}_{\Rightarrow}(\mathcal{E}^{(n)}), \quad (20)$$

$$R_{\Leftarrow} \leq \frac{1}{n} \bar{\chi}_{\Leftarrow}(\mathcal{E}^{(n)}). \quad (21)$$

Expressing $\bar{\chi}_{\Rightarrow}(\mathcal{E}^{(n)})$ as a telescopic sum and using the fact $\bar{\chi}_{\Rightarrow}(\mathcal{E}^{(0)}) = 0$,

$$\bar{\chi}_{\Rightarrow}(\mathcal{E}^{(n)}) = \sum_{t=1}^n [\bar{\chi}_{\Rightarrow}(\mathcal{E}^{(t)}) - \bar{\chi}_{\Rightarrow}(\mathcal{E}^{(t-1)})]. \quad (22)$$

$$= \sum_{t=1}^n \Delta \chi_{\Rightarrow}(\mathcal{E}^{(t-1)}), \quad (23)$$

where Eq. (23) comes from the fact that $\mathcal{E}^{(t)} = (\mathcal{A}_t \otimes \mathcal{B}_t) \mathcal{N} \mathcal{E}^{(t-1)}$, and $\mathcal{A}_t, \mathcal{B}_t$ leave $\bar{\chi}_{\Rightarrow}$ invariant so that $\bar{\chi}_{\Rightarrow}(\mathcal{E}^{(t)}) = \bar{\chi}_{\Rightarrow}(\mathcal{N} \mathcal{E}^{(t-1)})$. Likewise,

$$\bar{\chi}_{\Leftarrow}(\mathcal{E}^{(n)}) = \sum_{t=1}^n \Delta \chi_{\Leftarrow}(\mathcal{E}^{(t-1)}). \quad (24)$$

Putting Eqs. (20), (21), (23), and (24) together, we find

$$(R_{\Rightarrow}, R_{\Leftarrow}) \in \llbracket \frac{1}{n} \chi_{\Rightarrow}(\mathcal{E}^{(n)}), \frac{1}{n} \chi_{\Leftarrow}(\mathcal{E}^{(n)}) \rrbracket \quad (25)$$

$$\subset \text{conv}\{\llbracket \Delta \chi_{\Rightarrow}(\mathcal{E}), \Delta \chi_{\Leftarrow}(\mathcal{E}) \rrbracket : \text{arbitrary } \mathcal{E}\}, \quad (26)$$

which completes the proof. \square

As a side remark, we can also keep track of the entanglement E of the ensemble $\mathcal{E}^{(t)}$ along with the local Holevo information. We define $\Delta E(\mathcal{E}) := E(\mathcal{N} \mathcal{E}) - E(\mathcal{E})$, where $E(\{p_i, \rho_i\}) := \sum_i p_i E(\rho_i)$ and E is defined to be the distillable entanglement if $R_e \geq 0$ and the entanglement cost if $R_e \leq 0$. Thus, we have the following:

Theorem 3. If $(R_{\Rightarrow}, R_{\Leftarrow}, R_e)$ is achievable, then

$$(R_{\Rightarrow}, R_{\Leftarrow}, R_e) \in \text{conv}\{\llbracket \Delta \chi_{\Rightarrow}(\mathcal{E}), \Delta \chi_{\Leftarrow}(\mathcal{E}) \rrbracket \times [-\Delta E(\mathcal{E})^-, \Delta E(\mathcal{E})^+] : \text{arbitrary } \mathcal{E}\}. \quad (27)$$

However, in the remainder of Sec. 3, we will consider only the case of unlimited entanglement assistance.

3.2. Two-way communication protocols based on remote state preparation

Having derived outer bounds on the achievable region, we now turn to inner bounds based on protocols that achieve particular rates of communication. The main idea is to generalize the protocol from Sec. 4.3 of Ref. 21 (and see also generalizations in Refs. 22 and 29), originally designed for one-way communication with two-way

unitary channels, to the general problem of two-way communication with possibly nonunitary two-way channels. The main ingredients of the protocol of Ref. 21 are the HSW Theorem and *remote state preparation* (RSP), which allows Alice to share states of her choice with Bob. Our two-way protocol is based on two-way analogs of the HSW Theorem and RSP. In this section, we describe these tools, present the generalized two-way protocol, and analyze its error rate and inefficiency. In the following section, we show that combining the general technique with a particular method of remote state preparation gives an explicit inner bound.

First, we describe a bidirectional version of the HSW Theorem. Such a tool is necessary since the effective channel through which Alice can send signals to Bob depends on what input Bob is using to send signals to Alice, and vice versa. Consider any *two-way CQ channel*, i.e., a channel with two classical inputs, i for Alice and j for Bob, giving rise to a joint quantum output state ρ_{ij} . In this setting, we have the following:

Lemma 1 (Bidirectional HSW inner bound). *For the two-way CQ channel $\{i, j \rightarrow \rho_{i,j}\}$, rates $(R_{\Rightarrow}, R_{\Leftarrow})$ satisfying*

$$(R_{\Rightarrow}, R_{\Leftarrow}) \in \text{conv}\{\llbracket \bar{\chi}_{\Rightarrow}(\mathcal{E}), \bar{\chi}_{\Leftarrow}(\mathcal{E}) \rrbracket : \mathcal{E} = \{p_i q_j, \rho_{ij}\}\} \quad (28)$$

are achievable.

Proof (sketch). We omit the straightforward (but lengthy) generalization of the proof of Theorem 1 in Refs. 4 and 5. The basic idea is as follows. For each use of the two-way channel, Alice is unaware of Bob's input (which defines the effective channel from Alice to Bob), but such information *is* available for Bob in his decoding operation. The same holds for communication from Bob to Alice. Then, it is possible to show that good random codes (chosen independently by Alice and Bob) exist, and allow communication at the above rates according to a packing lemma analogous to that in the original proof. \square

Note that in this lemma, we have assumed that Alice and Bob choose their signals independently, with encoding distributions $\{p_i\}$ for Alice and $\{q_j\}$ for Bob, so that we can view Bob's encoding distribution as inducing a distribution over CQ channels from Alice to Bob, and vice versa. However, this is not the most general encoding distribution possible with many uses of the channel, so it is not clear whether this inner bound for CQ channels can be exceeded.

To prepare ensembles for bidirectional communication, we consider bipartite remote state preparation. Here the goal is to prepare a large number n of states drawn from the bipartite ensemble $\mathcal{E} = \{p_{ij}, \rho_{ij}\}$ with each party knowing one of the labels i, j . Note that Alice's label may control Bob's portion of the state as well as her own (indeed, the state may be entangled between their respective systems), and similarly for Bob's label.

We will assume the existence of a (not necessarily optimal) protocol for bipartite remote state preparation with known asymptotic classical communication and

entanglement costs $C_{\Rightarrow}, C_{\Leftarrow}$ and C_e . More specifically, suppose $n(C_{\Rightarrow} + \delta_n^{\text{RSP}})$ forward classical bits, $n(C_{\Leftarrow} + \delta_n^{\text{RSP}})$ backward classical bits, and $n(C_e + \delta_n^{\text{RSP}})$ ebits are sufficient for Alice and Bob to prepare a state drawn from $\mathcal{E}^{\otimes n}$ with fidelity $1 - \epsilon_n^{\text{RSP}}$, such that $\delta_n^{\text{RSP}}, \epsilon_n^{\text{RSP}} \rightarrow 0$ as $n \rightarrow \infty$. The problem of optimizing such costs in general is quite difficult, and has only been solved for very special cases of \mathcal{E} . But assuming the existence of such a protocol to prepare any particular ensemble of the form $\mathcal{E} = \{p_i q_j, \rho_{ij}\}$ (where the labels are chosen independently but the corresponding state may be arbitrary) at given costs, a corresponding point can be attained in the achievable region.

Lemma 2. $(R_{\Rightarrow}, R_{\Leftarrow}, -\infty)$ is achievable for all

$$\begin{aligned} (R_{\Rightarrow}, R_{\Leftarrow}) \in \text{conv}\{\llbracket \Gamma_{\Rightarrow}(\mathcal{E}), \Gamma_{\Leftarrow}(\mathcal{E}) \rrbracket : \mathcal{E} = \{p_a q_b, \rho_{ab}\} \text{ such that} \\ C_{\Rightarrow}(\mathcal{E}) \leq \bar{\chi}_{\Rightarrow}(\mathcal{N}\mathcal{E}) - \Gamma_{\Rightarrow}(\mathcal{E}), \\ C_{\Leftarrow}(\mathcal{E}) \leq \bar{\chi}_{\Leftarrow}(\mathcal{N}\mathcal{E}) - \Gamma_{\Leftarrow}(\mathcal{E})\}. \end{aligned} \quad (29)$$

Here $\bar{\chi}_{\Rightarrow}(\mathcal{N}\mathcal{E}), \bar{\chi}_{\Leftarrow}(\mathcal{N}\mathcal{E})$ represent achievable forward and backward communication rates for the ensemble $\mathcal{N}\mathcal{E}$ (according to Lemma 1). The quantities $\Gamma_{\Rightarrow}(\mathcal{E}), \Gamma_{\Leftarrow}(\mathcal{E})$ thus represent the amount of communication gained by one use of \mathcal{N} on the ensemble \mathcal{E} ; that is, the communication rates of $\mathcal{N}\mathcal{E}$ minus the communication costs of preparing \mathcal{E} .

Proof. By convexity and monotonicity, we only need to show that $R_{\Rightarrow} = \bar{\chi}_{\Rightarrow}(\mathcal{N}\mathcal{E}) - C_{\Rightarrow}(\mathcal{E})$ and $R_{\Leftarrow} = \bar{\chi}_{\Leftarrow}(\mathcal{N}\mathcal{E}) - C_{\Leftarrow}(\mathcal{E})$ are achievable rates (given a sufficiently large amount of entanglement assistance). We do this by giving a communication protocol achieving those rates assuming the existence of an RSP protocol with the stated communication costs. Since the protocol is a generalization of that in Sec. 4.3 of Ref. 21 for one-way communication, some readers may wish to refer to the detailed description and schematic diagram therein.

The protocol is as follows. Alice and Bob preagree on sufficiently large values of n and k (to be determined later) and proceed with the following protocol, using \mathcal{N} approximately nk times to communicate k messages a_1, a_2, \dots, a_k , each consisting of $n(R_{\Rightarrow} - \delta_n^{\text{HSW}} - \delta_n^{\text{RSP}})$ bits, in the forward direction, and k messages b_1, b_2, \dots, b_k , each consisting of $n(R_{\Leftarrow} - \delta_n^{\text{HSW}} - \delta_n^{\text{RSP}})$ bits, in the backward direction.

1. Using bipartite RSP, prepare a state ρ_1 from $\mathcal{E}^{\otimes n}$ with fidelity at least $1 - \epsilon_n^{\text{RSP}}$. This requires $O(n)$ uses of \mathcal{N} to communicate $n(C_{\Rightarrow}(\mathcal{E}) + \delta_n^{\text{RSP}})$ bits from Alice to Bob and $n(C_{\Leftarrow}(\mathcal{E}) + \delta_n^{\text{RSP}})$ bits from Bob to Alice. (Note that this initial communication is always possible if the appropriate $\Delta\chi$ is positive for some ensemble (and otherwise, it is not necessary). Suppose that although $\Delta\chi_{\Rightarrow} > 0$ for some ensemble, $\Delta\chi_{\Rightarrow} \leq 0$ for all ensembles that can be created at zero cost. In this case the operation is *semicausal* from Alice to Bob, and hence is also *semilocalizable*,^{30,31} meaning that it can be simulated by a local operation by Bob, sending a quantum state to Alice, and a final local operation by Alice.

But such an operation clearly has $\Delta_{\chi \Rightarrow} \leq 0$ for all ensembles, which is a contradiction. A similar argument applies to $\Delta_{\chi \Leftarrow}$.)

2. Apply $\mathcal{N}^{\otimes n}$ to ρ_1 , which has been chosen such that local measurements by Alice and Bob on $\mathcal{N}^{\otimes n}(\rho_1)$ provide an $n(\bar{\chi}_{\Rightarrow}(\mathcal{N}\mathcal{E}) - \delta_n^{\text{HSW}})$ -bit message for Bob and an $n(\bar{\chi}_{\Leftarrow}(\mathcal{N}\mathcal{E}) - \delta_n^{\text{HSW}})$ -bit message for Alice with probability at least $1 - \epsilon_n^{\text{RSP}} - \epsilon_n^{\text{HSW}}$, according to Lemma 1. Bob receives the message a_1 as well as the information needed to perform RSP of $\rho_2 \in \mathcal{E}^{\otimes n}$ in the next step. Similarly, Alice receives b_1 and together with the information she needs for RSP in the next round.
3. Perform the 2nd, \dots , k th rounds of RSP.

Just as in Ref. 21, Alice and Bob must know all of a_1, a_2, \dots, a_k and b_1, b_2, \dots, b_k at the beginning of the protocol, perform their measurements for the k th round of RSP to obtain the RSP instructions for the k th round, encode them as part of the message of the $(k-1)$ th round of RSP, and proceed with their part of the $(k-1)$ th round of RSP, and so on, until the first round RSP messages are generated and sent by the initial $O(n)$ uses of \mathcal{N} . For simplicity, we only consider non-interactive RSP protocols (with only one round of communication from Alice to Bob and vice versa), which will be sufficient for our applications. (With one more level of block coding, one should be able to use interactive RSP protocols, but this would require a more detailed error analysis.)

Finally, we analyze the errors and inefficiencies to show achievability of the rates. Fix any desired $\delta, \epsilon > 0$. The above protocol employs $n(c+k)$ uses of \mathcal{N} (for some constant c) and communicates $nk(\bar{\chi}_{\Rightarrow}(\mathcal{E}) - C_{\Rightarrow}(\mathcal{E}) - \delta_n^{\text{RSP}} - \delta_n^{\text{HSW}})$ and $nk(\bar{\chi}_{\Leftarrow}(\mathcal{E}) - C_{\Leftarrow}(\mathcal{E}) - \delta_n^{\text{RSP}} - \delta_n^{\text{HSW}})$ bits forward and backward, respectively, with error $k(\epsilon_n^{\text{HSW}} + \epsilon_n^{\text{RSP}})$. We can choose k, n independently large enough so that $\frac{\epsilon}{k} + \delta_n^{\text{RSP}} + \delta_n^{\text{HSW}} < \delta$ and then increase n if needed, to ensure $k(\epsilon_n^{\text{HSW}} + \epsilon_n^{\text{RSP}}) \leq \epsilon$. \square

3.3. Inner bound

Explicit inner bounds for the achievable region can be obtained from Lemma 2 together with known RSP protocols. In particular, we will make extensive use of the protocol for one-way RSP of entangled states:

Theorem 4 (RSP of entangled states³²). Asymptotically, the ensemble $\mathcal{E} = \{p_a, |\psi_a\rangle_{\text{AB}}\}$ can be prepared with a communication cost from Alice to Bob of $C_{\Rightarrow}(\mathcal{E}) = \chi_{\Rightarrow}(\mathcal{E})$ and a rate of entanglement consumption of $C_e(\mathcal{E}) = \sum_a p_a S(\text{tr}_A |\psi_a\rangle\langle\psi_a|)$.

Using this RSP protocol in the general inner bound of Lemma 2, we find

Theorem 5. $(R_{\Rightarrow}, R_{\Leftarrow}, -\infty)$ is achievable if

$$(R_{\Rightarrow}, R_{\Leftarrow}) \in \text{conv}\{[\Delta_{\chi \Rightarrow}(\mathcal{E}), \Delta_{\chi \Leftarrow}(\mathcal{E})] : \mathcal{E} = \{p_a, \rho_a\} \otimes \{q_b, \eta_b\}\} \quad (30)$$

In the ensemble of Eq. (30), the tensor product decomposition of AA'BB' may be arbitrary: Alice can prepare joint states of any subspace of her system and Bob's,

as can Bob, so long as the two subspaces are disjoint. In other words, we can have $\rho_a \in X, \eta_b \in Y$ for any fixed decomposition $AA'BB' = X \otimes Y$ into arbitrary complementary subspaces.

Proof. Alice and Bob are each given the knowledge of a and b . By Theorem 4, $C_{\Rightarrow}(\mathcal{E}) = \chi(\{p_a, \rho_a\})$ and $C_{\Leftarrow}(\mathcal{E}) = \chi(\{q_b, \eta_b\})$. Thus, the result follows from Lemma 2. \square

When one of the ensembles is trivial, the protocol performs one-way communication, with $\bar{\chi}_{\Leftarrow}(\mathcal{E}) = 0$ or $\bar{\chi}_{\Rightarrow}(\mathcal{E}) = 0$ as appropriate. For example, in the former case, the rate of forward communication $R_{\Rightarrow} = \Delta\chi_{\Rightarrow}(\mathcal{E})$ is achievable for arbitrary \mathcal{E} , so that the inner and outer bounds meet. Therefore, we find an expression for the entanglement-assisted one-way forward capacity of a two-way quantum channel. Similarly, we find $R_{\Leftarrow} = \Delta\chi_{\Leftarrow}(\mathcal{E})$ for the one-way backward capacity. Indeed, this result is immediate from the fact that the protocol of Ref. 21 for one-way communication applies unchanged even when \mathcal{N} is not unitary. Thus we have the following:

Corollary 1 (One-way capacity of a two-way channel).

$$R_{\Rightarrow}^{\max} := \sup\{R_{\Rightarrow} : (R_{\Rightarrow}, 0, -\infty) \text{ achievable}\} = \sup_{\mathcal{E}} \Delta\chi_{\Rightarrow}(\mathcal{E}) \quad (31)$$

$$R_{\Leftarrow}^{\max} := \sup\{R_{\Leftarrow} : (0, R_{\Leftarrow}, -\infty) \text{ achievable}\} = \sup_{\mathcal{E}} \Delta\chi_{\Leftarrow}(\mathcal{E}) \quad (32)$$

where the supremum is over all ensembles $\{p_i, \rho_{i,AA'BB'}\}$ with ancillary systems A', B' .

In particular, we have

Corollary 2. $R_{\Rightarrow, \Leftarrow}^{\max}$ are strongly additive. In other words, for any pair of two-way quantum channels $\mathcal{N}, \mathcal{N}'$, we have $R_{\Rightarrow}^{\max}(\mathcal{N} \otimes \mathcal{N}') = R_{\Rightarrow}^{\max}(\mathcal{N}) + R_{\Rightarrow}^{\max}(\mathcal{N}')$ and $R_{\Leftarrow}^{\max}(\mathcal{N} \otimes \mathcal{N}') = R_{\Leftarrow}^{\max}(\mathcal{N}) + R_{\Leftarrow}^{\max}(\mathcal{N}')$.

3.4. Relation to Shannon's classical bounds

Both the inner and outer bounds given above reduce to Shannon's bounds¹² in the case of a two-way classical channel (in which case entanglement assistance clearly does not help).

Consider sending information through a two-way classical channel. Suppose the input symbols a, b appear with the joint probability distribution p_{ab} . Then the output symbols a', b' appear with the joint probability distribution $p_{ab} p_{a'b'|ab}$, where the conditional probabilities $p_{a'b'|ab}$ define the channel. Let $I(X; Y|Z)$ denote the conditional mutual information,

$$I(X; Y|Z) := H(X|Z) - H(X|YZ), \quad (33)$$

where $H(X|Y)$ denotes the conditional Shannon entropy of X given Y . In terms of the conditional mutual information, Shannon proved the following inner and outer bounds on the capacity of a two-way classical channel:

Theorem 6 (Shannon¹²). If $(R_{\Rightarrow}, R_{\Leftarrow})$ is achievable, then

$$(R_{\Rightarrow}, R_{\Leftarrow}) \in \text{conv}\{\llbracket I(A; B'|B), I(B; A'|A) \rrbracket : p_{ab} \text{ arbitrary} \}. \quad (34)$$

Conversely, if

$$(R_{\Rightarrow}, R_{\Leftarrow}) \in \text{conv}\{\llbracket I(A; B'|B), I(B; A'|A) \rrbracket : p_{ab} = p_a q_b \}, \quad (35)$$

then $(R_{\Rightarrow}, R_{\Leftarrow})$ is achievable.

Now consider the corresponding outer and inner bounds from Theorems 2 and 5. Let $\mathcal{E} = \{p_{ab}, |aa\rangle_{AA'}|bb\rangle_{BB'}\}$ where $|a\rangle, |b\rangle$ are mutually orthogonal states on systems A, B and as before, the senders retain copies of their inputs in systems A', B' . The action of the two-way classical channel \mathcal{N} on this ensemble is

$$\mathcal{N}(|a\rangle\langle a| \otimes |b\rangle\langle b|) = \sum_{a', b'} p_{a'b'|ab} |a'\rangle\langle a'| \otimes |b'\rangle\langle b'|. \quad (36)$$

It is straightforward to compute

$$\chi(\text{tr}_{AA'} \mathcal{E}) = H(B) \quad (37)$$

$$\chi(\text{tr}_{AA'} \mathcal{N}\mathcal{E}) = H(BB') - H(BB'|AB) = H(B) + I(B'; A|B). \quad (38)$$

Therefore,

$$\Delta\chi_{\Rightarrow}(\mathcal{E}) = I(A; B'|B). \quad (39)$$

Thus, we see that in the classical case, the outer bound of Theorem 2 and the inner bound of Theorem 5 are identical to Shannon's outer and inner bounds, respectively.

The equivalence to Shannon's bounds shows that in general, the bounds of Theorems 2 and 5 are not tight; even in the classical case, the inner bound may be exceeded^{14, 16–18} and the outer bound may not be achievable.^{19, 20} Such results for classical channels might provide insight into how the general (quantum) bounds could be tightened.

4. Additivity results for one-way channels

Since one-way quantum channels are simply special cases of two-way channels, it is possible to obtain results about one-way channels by thinking of them as two-way channels. In this section, we use such an approach to rederive two previously known additivity results for one-way channels: the entanglement-assisted capacity of an arbitrary one-way quantum channel, and the entanglement-unassisted capacity of an entanglement-breaking one-way quantum channel.

In this section, \mathcal{M} denotes a one-way channel, and the classical capacity is simply the maximum value of R_{\Rightarrow} in the achievable region (at fixed R_e). We use $R(\mathcal{M})$ and $R^E(\mathcal{M})$ to denote the classical capacity with no entanglement assistance and unlimited entanglement assistance, respectively.

4.1. Entanglement-assisted capacity of one-way channels

A general expression for $R^E(\mathcal{M})$ in terms of the quantum mutual information was found in Ref. 7. Furthermore, Ref. 7 proved that R^E is strongly additive, i.e., $R^E(\mathcal{M}_1 \otimes \mathcal{M}_2) = R^E(\mathcal{M}_1) + R^E(\mathcal{M}_2)$ for any pair of (one-way) quantum channels $\mathcal{M}_1, \mathcal{M}_2$.

The original proof of additivity used entropy inequalities to show that the explicit expression for R^E is indeed additive.^{7,33} But specializing Corollary 2 to one-way quantum channels provides an immediate alternative proof of strong additivity.

These two proofs appear to be inequivalent. The simplicity of proving additivity via Corollary 2 seems to follow from the structure of the protocol of Ref. 21 (or equivalently, that in Lemma 2). The main idea of this protocol, to borrow a resource and later regenerate some or more of it, has recently found a number of applications in quantum information theory.^{22,23,34} Such a protocol gives rise to a coding structure very different from more standard, direct techniques, such as those used in Ref. 7.

From Corollary 1, we see that the capacity expression of Ref. 7 in terms of the quantum mutual information can be written as a supremum of $\Delta\chi$. It is not obvious simply by looking at these two expressions that they are in fact equal.

4.2. Unassisted capacity of one-way entanglement-breaking channels

We now turn our attention to unassisted classical communication using a one-way channel. In particular, we consider *entanglement-breaking channels*, which are guaranteed to output a state that is unentangled between the sender and the receiver. Using the framework of two-way channels, we will prove a special case of the following result:

Theorem 7 (Shor²⁷). If \mathcal{M} is an arbitrary one-way quantum channel and \mathcal{M}' is an entanglement-breaking one-way quantum channel, then

$$R(\mathcal{M} \otimes \mathcal{M}') = R(\mathcal{M}) + R(\mathcal{M}'). \quad (40)$$

We will prove this result in the special case in which both \mathcal{M} and \mathcal{M}' are entanglement-breaking. In particular, this includes the case $\mathcal{M} = \mathcal{M}'$, demonstrating the additivity of the Holevo capacity of an entanglement-breaking channel. The proof in terms of two-way channels for this special case is significantly simpler.

As in Shor's proof,²⁷ we use strong subadditivity of the von Neumann entropy³⁵ in various guises. In particular, we will use the following lemma:

Lemma 3. Let $\{\sigma_i\}, \{\eta_i\}$ be sets of quantum states, and let $\{p_i\}$ be a probability distribution. Then

$$S(\sum_i p_i \sigma_i \otimes \eta_i) \geq S(\sum_i p_i \sigma_i) + \sum_i p_i S(\eta_i). \quad (41)$$

We give two proofs of this lemma: an operational proof and a proof that uses strong subadditivity directly.

Proof 1 (Operational). Let $\mathcal{E}_1 = \{p_i, \sigma_i\}$ and $\mathcal{E}_2 = \{p_i, \sigma_i \otimes \eta_i\}$. Since \mathcal{E}_1 can be obtained from \mathcal{E}_2 by discarding the second system,

$$0 \leq \chi(\mathcal{E}_2) - \chi(\mathcal{E}_1) \quad (42)$$

$$= S(\sum_i p_i \sigma_i \otimes \eta_i) - \sum_i p_i S(\eta_i) - S(\sum_i p_i \sigma_i), \quad (43)$$

where the last line is obtained by using the definition Eq. (9) and the fact that $S(\sigma \otimes \eta) = S(\sigma) + S(\eta)$. \square

Proof 2 (Direct use of strong subadditivity). For the state $\rho_{ABC} := \sum_i p_i \sigma_{i,A} \otimes \eta_{i,B} \otimes |i\rangle\langle i|_C$,

$$S_{ABC} = H(\{p_i\}) + \sum_i p_i S(\sigma_i \otimes \eta_i) \quad (44)$$

$$S_{AB} = S(\sum_i p_i \sigma_i \otimes \eta_i) \quad (45)$$

$$S_{AC} = H(\{p_i\}) + \sum_i p_i S(\sigma_i) \quad (46)$$

$$S_A = S(\sum_i p_i \sigma_i). \quad (47)$$

Equation (41) then follows from the strong subadditivity inequality $S_{ABC} + S_A \leq S_{AB} + S_{AC}$. \square

Proof of Theorem 7 for \mathcal{M} entanglement-breaking. The idea of the proof is to show that the states that can be output by either channel are of no use in enhancing the capacity of the other channel, and hence that the capacity of the joint channel is simply the sum of the individual capacities.

For any entanglement-unassisted protocol that uses only entanglement-breaking channels, we can rerun our proof of the outer bound in Sec. 3.1 restricting to ensembles of separable states. Thus we have an upper bound analogous to Eq. (24),

$$R(\mathcal{M}) \leq \max\{\Delta\chi_{\Rightarrow}(\mathcal{E}) : \text{separable } \mathcal{E}\}. \quad (48)$$

Applying the two-way channel formalism to a one-way channel \mathcal{M} with input system A and output system B, the most general input and output ensembles are $\mathcal{E}_{\text{in}} = \{p_i, \rho_{i,A'AB'}\}$ and $\mathcal{E}_{\text{out}} = \{p_i, \mathcal{M}(\rho_{i,A'AB'})\}$. Without loss of generality, we can omit the system A'. This system does not appear in the bound on the communication rate in terms of $\Delta\chi_{\Rightarrow}$, and there is no entanglement to be stored in A'. Thus, the optimal input ensemble can be restricted to have the form $\mathcal{E}_{\text{in}} = \{p_i, \rho_{i,AB'}\}$ with $\mathcal{E}_{\text{out}} = \{p_i, \mathcal{M}(\rho_{i,AB'})\}$.

We will be interested in three cases where the form of the input ensemble is restricted to different extents. Let $\Delta\chi^S$, $\Delta\chi^P$, $\Delta\chi^0$ denote $\sup_{\mathcal{E}_{\text{in}}} [\chi(\mathcal{E}_{\text{out}}) - \chi(\text{tr}_A \mathcal{E}_{\text{in}})]$ for \mathcal{E}_{in} ranging over

$$\mathcal{E}_{\text{in}}^S := \{p_i, \sum_j q_{ij} \rho_{ij,A} \otimes \sigma_{ij,B'}\}, \quad (49)$$

$$\mathcal{E}_{\text{in}}^P := \{p_i, \rho_{i,A} \otimes \sigma_{i,B'}\}, \quad (50)$$

$$\mathcal{E}_{\text{in}}^0 := \{p_i, \rho_{i,A}\} \quad (51)$$

for separable ensembles, product ensembles, and ensembles with $\chi = 0$, respectively. It is clear that $\Delta\chi^0 \leq \Delta\chi^P \leq \Delta\chi^S$; we will show that $\Delta\chi^S \leq \Delta\chi^P \leq \Delta\chi^0$, so that in fact, all three quantities are equal.

First we show that $\Delta\chi^S \leq \Delta\chi^P$. For any separable ensemble in the form of Eq. (49), let $\eta_{ij,A} := \mathcal{M}(\rho_{ij,A})$. Then we have

$$\begin{aligned} \chi(\mathcal{E}_{\text{out}}) - \chi(\text{tr}_A \mathcal{E}_{\text{in}}) &= S(\sum_{ij} p_i q_{ij} \sigma_{ij,B'} \otimes \eta_{ij,A}) - \sum_i p_i S(\sum_j q_{ij} \sigma_{ij,B'} \otimes \eta_{ij,A}) \\ &\quad - S(\sum_{ij} p_i q_{ij} \sigma_{ij,B'}) + \sum_i p_i S(\sum_j q_{ij} \sigma_{ij,B'}) \end{aligned} \quad (52)$$

$$\begin{aligned} &\leq S(\sum_{ij} p_i q_{ij} \sigma_{ij,B'} \otimes \eta_{ij,A}) - \sum_{ij} p_i q_{ij} S(\sigma_{ij,B'} \otimes \eta_{ij,A}) \\ &\quad - S(\sum_{ij} p_i q_{ij} \sigma_{ij,B'}) + \sum_{ij} p_i q_{ij} S(\sigma_{ij,B'}) \leq \Delta\chi^P \end{aligned} \quad (53)$$

where Lemma 3 has been applied to each term in $\sum_i p_i S(\cdot)$. Thus, a separable ensemble is no better than a product ensemble.

Now we show that $\Delta\chi^P \leq \Delta\chi^0$. For any product ensemble in the form of Eq. (50), let $\eta_{i,A} := \mathcal{M}(\rho_{i,A})$. Then

$$\begin{aligned} \chi(\mathcal{E}_{\text{out}}) - \chi(\text{tr}_A \mathcal{E}_{\text{in}}) &= S(\sum_i p_i \sigma_{i,B'} \otimes \eta_{i,A}) - \sum_i p_i S(\sigma_{i,B'} \otimes \eta_{i,A}) \\ &\quad - S(\sum_i p_i \sigma_{i,B'}) + \sum_i p_i S(\sigma_{i,B'}) \end{aligned} \quad (54)$$

$$\leq S(\sum_i p_i \eta_{i,A}) - \sum_i p_i S(\eta_{i,A}) \leq \Delta\chi^0 \quad (55)$$

where the inequality is due to subadditivity of $S(\sum_i p_i \sigma_i \otimes \eta_i)$ and additivity of $S(\sigma \otimes \eta)$. Thus a product ensemble with $\chi \neq 0$ is no better than one with $\chi = 0$.

This argument shows that we can assume without loss of generality that the input ensemble is of the form of Eq. (51). But such an ensemble costs nothing to create, so by using the protocol of Lemma 2, we see that the capacity $\Delta\chi^0$ can be achieved for any ensemble $\mathcal{E}_{\text{in}}^0$, and

$$R(\mathcal{M}) = \sup_{\mathcal{E}_{\text{in}}^0} \Delta\chi^0. \quad (56)$$

Finally, consider the capacity of the combined channel $\mathcal{M} \otimes \mathcal{M}'$ where both \mathcal{M} and \mathcal{M}' are entanglement-breaking. Without loss of generality, we can assume that the channels act sequentially. Each channel can only produce separable output states, which by the above argument are no better than states with $\chi = 0$, which can be produced at zero cost. Therefore the capacity of the combined channel is simply the sum of the individual capacities. \square

5. Open questions

In this paper, we have established simple inner and outer bounds on the entanglement-assisted classical capacity region of a two-way quantum channel, and we have applied the framework of two-way channels to rederive two previous additivity results for one-way channels. However, since the two-way channel framework includes a wide variety of disparate communication scenarios as special cases, this work raises many more questions than it answers.

Calculating the capacity region for any particular channel can be a challenging problem, and has not been done except in a few particular special cases. In fact, even computing the inner and outer bounds given in this paper can be difficult, since the ancillary state spaces A', B' may be arbitrarily large, and we do not know that low-dimensional ancillas are sufficient to achieve the capacity (even in the unitary case²¹).

Although calculating the precise capacity region may be difficult, a more modest goal is to improve upon the inner and outer bounds given in this paper. In particular, known classical bounds that improve upon Shannon's bounds^{14, 16–20} might be useful for finding improved quantum bounds. Also, it would be interesting to find conditions under which the inner and outer bounds coincide. Shannon showed that his inner and outer bounds coincide for certain kinds of symmetric two-way classical channels,¹² so it is plausible to suppose that a similar result might hold in the quantum case.

We have primarily considered the case of unlimited entanglement assistance, but it would be interesting to consider unassisted communication as well as the general case of finite entanglement assistance. Recently, Shor has given a protocol for classical communication through a one-way quantum channel with limited entanglement assistance that interpolates between the HSW capacity and the entanglement-assisted capacity.³⁸ In addition, Harrow has obtained an expression for the one-way classical communication capacity with finite entanglement assistance for unitary two-way channels.²² It would be interesting to generalize these results to arbitrary two-way channels.

Another approach is to consider particular families of channels to see whether the capacity region is simpler for those channels. One such family is the set of two-way entanglement-breaking channels. There are several possible definitions of a two-way entanglement-breaking channel, but perhaps the simplest is that the output state should be triseparable between Alice's output, Bob's output, and any ancillas. Unfortunately, it is not even clear how to characterize such channels (as can be done for one-way entanglement-breaking channels^{36, 37}). Some results have been obtained for other families of two-way quantum channels, such as unitary two-way channels^{21–23} and feedback channels.²⁶ Another special class of two-way channels, those that simply distribute bipartite states, have been much better understood.³⁴

One way to obtain a better understanding of channel capacities is to consider the problem of simulating a channel using a certain amount of communication (in each direction) and entanglement. Such *reverse theorems* have been studied for one-way classical⁷ and quantum³⁹ channels, and more recently for feedback channels with restricted input sources.²⁶ In particular, reverse theorems can be useful for establishing bounds on capacities.²¹ However, simple reverse theorems for general two-way channels seem unlikely to exist. For example, the communication costs (in each direction) of any simulation must exceed the corresponding one-way capacities, since the simulated channel can be used to achieve the one-way capacity in either direction. As another example, the set of causal operations that are not

localizable³⁰ cannot produce any communication, but any such operation requires communication in at least one direction to simulate, even with entanglement assistance.

Finally, note that we have completely avoided the problems of communicating quantum information through a two-way quantum channel and of multi-way communication through k -partite quantum operations with $k > 2$. These problems present further challenges for understanding the capabilities of quantum communication channels.

Acknowledgments

We thank Aram Harrow for many helpful discussions, especially regarding the protocol for achieving the entanglement-assisted one-way capacity of a two-way channel. Thanks also to Charles Bennett, whose insights on entanglement-assisted communication via entanglement breaking channels and the simulation of nonlocal boxes saved us from wandering in unfruitful directions. This work was initiated while AMC and DWL were at the IBM T. J. Watson Research Center and HKL was at Magiq Technologies. Part of this work was done while AMC was at the MIT Center for Theoretical Physics and while DWL was a visiting researcher at MSRI. AMC was supported in part by the Fannie and John Hertz Foundation, the Cambridge-MIT Foundation, the DOE under cooperative research agreement DE-FC02-94ER40818, and the NSA and ARDA under ARO contract DAAD19-01-1-0656. AMC and DWL received support from the NSF under Grant No. EIA-0086038. DWL received support from the Tolman Foundation and the Croucher Foundation. HKL received support from NSERC, the CRC Program, CFI, OIT, PREA, and CIPI.

References

1. C. E. Shannon, *A mathematical theory of communication*, Bell System Tech. J. **27**, 379 (1948).
2. A. S. Holevo, *Bounds for the quantity of information transmitted by a quantum communication channel*, Problems of Information Transmission **9**, 177 (1973).
3. P. Hausladen, R. Jozsa, B. Schumacher, M. Westmoreland, and W. K. Wootters, *Classical information capacity of a quantum channel*, Phys. Rev. A **54**, 1869 (1996).
4. B. Schumacher and M. D. Westmoreland, *Sending classical information via noisy quantum channels*, Phys. Rev. A **56**, 131 (1997).
5. A. S. Holevo, *The capacity of the quantum channel with general signal states*, IEEE Trans. Inf. Theory **44**, 269 (1998), quant-ph/9611023.
6. C. H. Bennett, P. W. Shor, J. A. Smolin, and A. V. Thapliyal, *Entanglement-assisted classical capacity of noisy quantum channels*, Phys. Rev. Lett. **83**, 3081 (1999), quant-ph/9904023.
7. C. H. Bennett, P. W. Shor, J. A. Smolin, and A. V. Thapliyal, *Entanglement-assisted capacity of a quantum channel and the reverse Shannon theorem*, IEEE Trans. Inf. Theory **48**, 2637 (2002), quant-ph/0106052.
8. S. Lloyd, *Capacity of the noisy quantum channel*, Phys. Rev. A **55**, 1613 (1997), quant-ph/9604015.

9. H. Barnum, M. A. Nielsen, and B. Schumacher, *Information transmission through a noisy quantum channel*, Phys. Rev. A **57**, 4153 (1998), quant-ph/9702049.
10. P. W. Shor, *The quantum channel capacity and coherent information*, MSRI Workshop on Quantum Computation, 2002, <http://www.msri.org/publications/ln/msri/2002/quantumcrypto/shor/1/>.
11. I. Devetak, *The private classical capacity and quantum capacity of a quantum channel*, IEEE Trans. Inf. Theory **51**, 44 (2005), quant-ph/0304127.
12. C. E. Shannon, *Two-way communication channels*, Proc. 4th Berkeley Symp. Math. Stat. Prob. (UC Press, Berkeley, 1961).
13. E. C. van der Meulen, *A survey of multi-way channels in information theory: 1961–1976*, IEEE Trans. Inf. Theory **23**, 1 (1977).
14. G. Dueck, *The capacity region of the two-way channel can exceed the inner bound*, Inform. Contr. **40**, 258 (1979).
15. A. El Gamal and T. M. Cover, *Multiple user information theory*, Proc. IEEE **68**, 1466 (1980).
16. J. P. M. Schalkwijk, *The binary multiplying channel—a coding scheme that operates beyond Shannon’s inner bound region*, IEEE Trans. Inf. Theory **28**, 107 (1982).
17. J. P. M. Schalkwijk, *On an extension of an achievable rate region for the binary multiplying channel*, IEEE Trans. Inf. Theory **29**, 445 (1983).
18. T. S. Han, *A general coding scheme for the two-way channel*, IEEE Trans. Inf. Theory **30**, 35 (1984).
19. Z. Zhang, T. Berger, and J. P. M. Schalkwijk, *New outer bounds to capacity regions of two-way channels*, IEEE Trans. Inf. Theory **32**, 383 (1986).
20. A. P. Hekstra and F. M. J. Willems, *Dependence balance bounds for single-output two-way channels*, IEEE Trans. Inf. Theory **35**, 44 (1989).
21. C. H. Bennett, A. Harrow, D. W. Leung, and J. A. Smolin, *On the capacities of bipartite Hamiltonians and unitary gates*, IEEE Trans. Inf. Theory **49**, 1895 (2003), quant-ph/0205057.
22. A. W. Harrow, *Coherent communication of classical messages*, Phys. Rev. Lett. **92**, 097902 (2004), quant-ph/0307091.
23. A. W. Harrow and D. W. Leung, *Bidirectional coherent classical communication*, quant-ph/0412126.
24. M. S. Leifer, L. Henderson, and N. Linden, *Optimal entanglement generation from quantum operations*, Phys. Rev. A **67**, 012306 (2003), quant-ph/0205055.
25. D. W. Berry and B. C. Sanders, *Relation between classical communication capacity and entanglement capability for two-qubit unitary operations*, Phys. Rev. A **68**, 032312 (2003), quant-ph/0207065.
26. I. Devetak, *A triangle of dualities: reversibly decomposable quantum channels, source-channel duality, and time reversal*, quant-ph/0505138.
27. P. W. Shor, *Additivity of the classical capacity of entanglement-breaking quantum channels*, J. Math. Phys. **43**, 4334 (2002), quant-ph/0201149.
28. M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, U.K., 2000).
29. A. W. Harrow, *Applications of Coherent Classical Communication and the Schur Transform in Quantum Information Theory*, Ph.D. thesis, Massachusetts Institute of Technology (2005).
30. D. Beckman, D. Gottesman, M. A. Nielsen, and J. Preskill, *Causal and localizable quantum operations*, Phys. Rev. A **64**, 052309 (2001), quant-ph/0102043.
31. T. Eggeling, D. Schlingemann, and R. F. Werner, *Semicausal operations are semilocalizable*, Europhys. Lett. **57**, 782 (2002), quant-ph/0104027.

32. C. H. Bennett, P. Hayden, D. W. Leung, P. W. Shor, and A. Winter, *Remote preparation of quantum states*, IEEE Trans. Inf. Theory **51**, 56–74 (2005), quant-ph/0307100.
33. N. Cerf and C. Adami, *Von Neumann capacity of noisy quantum channels*, Phys. Rev. A **56**, 3470 (1996), quant-ph/9609024.
34. I. Devetak, A. W. Harrow, and A. Winter, *A family of quantum protocols*, Phys. Rev. Lett. **93**, 230504 (2004), quant-ph/0308044.
35. E. Lieb and M. B. Ruskai, *Proof of the strong subadditivity of quantum mechanical entropy*, J. Math. Phys. **14**, 1938 (1973).
36. M. B. Ruskai, *Qubit entanglement breaking channels*, Rev. Math. Phys. **15**, 643 (2003), quant-ph/0302032.
37. M. Horodecki, P. W. Shor, and M. B. Ruskai, *General entanglement breaking channels*, Rev. Math. Phys. **15**, 629 (2003), quant-ph/0302031.
38. P. W. Shor, *The classical capacity achievable by a quantum channel assisted by limited entanglement*, in *Quantum Information, Statistics, Probability*, ed. Osamu Hirota (Rinton Press, Princeton, 2004), p. 144, quant-ph/0402129.
39. C. H. Bennett, I. Devetak, A. Harrow, P. W. Shor, and A. Winter, *The quantum reverse Shannon theorem*, in preparation.